INFORMATION AND CONTROL 69, 12–40 (1986)

On the Complexity of Inductive Inference*

ROBERT P. DALEY

Department of Computer Science, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

AND

CARL H. SMITH

Department of Computer Science, University of Maryland, College Park, Maryland 20742

The notion of the complexity of performing an inductive inference is defined. Some examples of the tradeoffs between the complexity of performing an inference and the accuracy of the inferred result are presented. An axiomatization of the notion of the complexity of inductive inference is developed and several results are presented which both resemble and contrast with results obtainable for the axiomatic computational complexity of recursive functions. © 1986 Academic Press, Inc.

1. INTRODUCTION

Inductive inference has been an important subject of investigation in the philosophy of science (see Hempel, 1965) for quite some time, where the focus has been on the development of a methodology (i.e., an algorithm) which would, given sufficient observations of some phenomenon, arrive at a correct theory for that phenomenon. By its nature inductive inference is a form of learning in the limit (i.e., in the long run), since until sufficient observations of the right sort have been made no means of distinguishing erroneous theories from correct theories based upon observation is available. With the development of mathematical logic beginning with Frege in the last century the formal tools became available with which the problem could be rigorously analyzed, with the result that the majority of philosophers of science are convinced that in general no such algorithm could exist. However, within the past decade inductive inference has

* Supported by NSF Grants MCS 7803617, MCS 8017332, MCS 7903912, MCS 8105214, and MCS 8301536. These results were presented at the Mathematical Foundations of Computer Science Conference in Prague, Czechoslovakia in September 1984.

0019-9958/86 $3.00

become an important area of investigation by computer scientists, for whom the main interest is the development and study of algorithms, not for the general case, but for specific classes of problems (i.e., heuristics for performing inferences). As such, these studies form a theoretical foundation for artificial intelligence. Instead of algorithms for inferring scientific theories given observational data, researchers in this area study algorithms which can infer a program (i.e., theory) for a computable function (i.e., phemonenon) given examples of the input/output behavior of the function (i.e., data). Other studies have dealt with algorithms which infer grammars given samples of the words which they generate. Several recent survey papers (see Angluin and Smith, 1983; Klette and Wiehagen, 1980), detail the wide range of investigations and their most important results.

The abstract study of inductive inference has focused on distinguishing various criteria for successful inference by a given class of machines. Herein a notion of the complexity of the inference process is presented with some examples of trade-offs between the complexity of an inference and the accuracy of the result of the inference. Then an axiomatization of the concept of the complexity of inductive inference is introduced. Our axiomatization parallels the approach made by Blum (1967) for the complexity of computations. Earlier studies of the complexity of inference were concerned with showing that certain inference problems were members of well-known complexity classes like P and NP (Angluin, 1978; Gold, 1978; Pudlak, 1975; Pudlak and Springsteel, 1975). Freivalds (1975) also attempted an axiomatization of a notion of complexity of inductive inference. The basic approach employed below is similar to his on a number of points. He was particularly concerned with incorporating into his axiomatization the number of distinct hypotheses made by an inference strategy as a possible measure of the complexity of inference. The number of distinct hypotheses has been used as a measure of the complexity of inference by some authors (see Barzdin and Freivalds, 1972). Another commonly used measure of the complexity of inference is the number of mind changes, and trade-offs between mind changes and accuracy have been studied in (Case and Ngo-Manguelle, 1979; Case and Smith, 1983). Recently and independently of this work, Shäfer-Richter (1984) has also provided an axiomatization of the complexity of inductive inference, which is different from ours, but which is a viable alternative. Our approach to the complexity of inductive inference will be based on the total resources used by an inference strategy in the process of converging to a correct inference. If one were to plot the resources used by an inference strategy for each new datum, then our notion of inference complexity could be represented as the area under this curve from the beginning of the inference process until the process converged to the correct explanation. By providing an axiomatization of the notion of the complexity of inference we hope to reveal those properties which must be shared by all measures of the complexity of inductive inference.

The remainder of this section is devoted to several notational conventions and preliminary definitions. The set of natural numbers is denoted by N. We use $\subseteq$, $\subset$ to denote respectively containment and proper containment for sets (including sets of ordered pairs), respectively. The class of all total recursive functions is denoted by $\mathscr{R}$, and $\mathscr{S}$ will range over subsets of $\mathscr{R}$. The function symbols f, g, h, r will range over $\mathscr{R}$, and ϕ will range over partial recursive functions. Also, α and β denote program transformations, i.e., total recursive funtions whose inputs are programs. We use $f(x)\downarrow$ and $f(x)\uparrow$ to indicate that $f(x)$ is defined and $f(x)$ is undefined respectively. The domain and range of the function f is denoted by dom f and ran f respectively. By ran $f \leqslant n$ we will mean that $(\forall x \in \text{dom } f)[f(x) \leqslant n]$. We use $f(x)\downarrow \neq g(x)$ to mean $f(x)\downarrow$ and $f(x) \neq g(x)$. Also, we use $f \mid n$ to denote the finite initial segment of f whose domain consists of $\{x \mid x \leqslant n\}$. We use σ to denote a finite function whose domain is an initial segment of N, and $\max_\sigma$ to denote $\max\{x \in \text{dom } \sigma\}$. The largest proper initial segment of σ, $\sigma \mid \max_\sigma - 1$, will be abbreviated as $\sigma - 1$. We write $\sigma \subseteq \sigma'$ to denote that $(\forall x \in \text{dom } \sigma)[\sigma'(x) = \sigma(x)]$. For any $h \in \mathscr{R}$, h^0 is the identity function and $h^{n+1} = h(h^n)$. We also use *0* to denote the constant function whose value is 0. If v is an n-tuple of values then π_k will denote the kth projection. The cardinality of the set X is denoted by card(X). We use $\phi_1 =^n \phi_2$ (read: ϕ_1 is an *n-variant* of ϕ_2) to mean that $\text{card}(\{x \mid \phi_1(x) \neq \phi_2(x)\}) \leqslant n$, and $\phi_1 =^* \phi_2$ to mean that $\{x \mid \phi_1(x) \neq \phi_2(x)\}$ is finite. The empty set is denoted by $\emptyset$, where $\max(\emptyset) = 0$. The quantifiers $\exists^\infty$ and $\forall^\infty$ stand for "there exist infinitely many" and "for all but finitely many," respectively. The sequence $\{\phi_i\}$ denotes an arbitrary *acceptable programming system* (see Machtey and Young, 1978), also known as an *acceptable numbering* of all and only the partial recursive functions (Rogers, 1958; Rogers, 1967).

Inferences will be performed by *Inductive Inference Machines* (IIMs) as defined recursion theoretically in (Blum and Blum, 1975), and used in essence previously in (Gold, 1967; Pulnam, 1975). An IIM M operates continuously as follows: M is presented with successively larger segments of the graph of some function f and in response produces a sequence p_1, p_2,..., of hypothesized programs for f. We say that M EX^n identifies f (and write $f \in EX^n(M)$) if there is some program p such that $\phi_p =^n f$ and p is the last *distinct* program produced by M, i.e., either M produces a finite nonempty sequence of programs and p is the last one, or M produces an infinite sequence of programs all but finitely many of which are p. Observe that M's sequence of hypotheses converges when either the sequence is eventually constant or is finite and nonempty. Thus, M can be undefined on all but a finite number of the initial segments of a function f and still converge to an acceptable program for f. This definition is easily seen to be equivalent to similar definitions, for example, (Blum and Blum, 1975; Case and Smith, 1983). The power of an IIM M is given by the set of functions which it can successfully infer. The class of sets of functions

$$EX^n = \{\mathscr{S} \mid (\exists M)[\mathscr{S} \subseteq EX^n(M)]\},$$

consists of all the sets of functions which can be successfully inferred by some IIM with respect to EX^n type inference. Our attention will be restricted to sets of total recursive functions (i.e., subsets of $\mathscr{R}$), and our interest will lie in inferring programs for total recursive functions only. There are several assumptions which can be made about IIMs without loss of generality with respect to inferrability. For example, M can be assumed to be total, i.e., defined on any finite segment of the graph of any function, since any partial IIM can be replaced by another IIM which is total and which infers the same functions as M (and perhaps more). Also, it can be assumed that the function f is presented to M (or querried by M) in increasing order with respect to its domain (i.e., f is presented as $f \mid 0$, $f \mid 1$, $f \mid 2$,...), since M after responding to $f \mid n$ could simply wait (repeating the response to $f \mid n$) until receiving $f \mid n+1$ before responding with any new hypothesis. Although such assumptions do not effect inferrability, they can alter the complexity of some inferences.

Technically, one would expect that the order of presentation of data to an IIM would affect the complexity of the inference. Gold (1967) found that a judicious choice of input orders can even increase the scope or power of an IIM. For the most part, however, the IIMs considered below will assume that functions are presented to them in increasing domain order. In the last section of this paper some examples of the effect of changing the order of presentation on the complexity of inference are given. Despite the effect of order on complexity, the restriction to a single order of presentation (increasing domain order being simply the most natural) can be justified on the grounds that inference is really a mass problem—an IIM must attempt to infer programs for many functions, and the real difficulty of doing so lies in discriminating the input function from the mass of possible functions. It is unrealistic to imagine that an IIM would perform experiments on (or encounter) each function in its most pernicious (or efficacious) order of presentation. Indeed, the increase or reduction in the complexity of inference by such a contrived presentation would probably be more than offset by the complexity involved in producing such a presentation. In fact, Young (1971), used priority techniques to produce faster enumerations of some sets by changing the order of the output. Perhaps our assumption reflects an idea of Hempel (1965), that in setting up an experiment, in deciding which data to gather, already some inferential effort has taken place.

Each IIM M will be chosen from an acceptable numbering and, hence, will be an effective device. Consequently, $\phi_M(f \mid n)$ will be used to denote the output value (if it exists) of M given the initial segment $f \mid n$ of the graph of f as input. We will use $\phi_M(f)$ to denote the limit (if it exists) $\lim_{n \to \infty} \phi_M(f \mid n)$. This limit exists if and only if for some m, $\phi_M(f \mid m)\downarrow$ and for all $n > m$ either $\phi_M(f \mid n)\uparrow$ or $\phi_M(f \mid n) = \phi_M(f \mid m)$. The least m satisfying these conditions represents the point of convergence of ϕ_M on the input function f (denoted by $\mu_M(f)$) and will play a crucial role in our analysis of the complexity of inductive inference. We view the input $f \mid n$ as being encoded as an integer, so that ϕ_M is simply a partial recursive function with certain limit properties, viz., $\phi_M(f)$ is the limit of ϕ_M on the sequence of integers $f \mid 0$, $f \mid 1$, $f \mid 2$,.... If one regards f as the limit of its finite initial segments $f \mid n$, which in turn may be regarded as approximations to f, then this view expresses a form of continuity where $\phi_M(f \mid n)$ is regarded as an approximation to $\phi_M(f)$. This is the basis of our use of the notation $\phi_M(f) = \lim_{n \to \infty} \phi_M(f \mid n)$.

We now formalize all of the preceding. For each partial recursive function ϕ and each $f \in \mathscr{R}$ we define,

$$\lim_{n \to \infty} \phi(f \mid n) = \begin{cases} \phi(f \mid m), & \text{where } m \text{ is the least integer such that} \\ & \phi(f \mid m)\downarrow \text{ and for all } n > m \text{ either} \\ & \phi(f \mid n)\downarrow = \phi(f \mid m) \text{ or } \phi(f \mid n)\uparrow, \\ \uparrow, & \text{if no such } m \text{ exists.} \end{cases}$$

A functional $\psi: \mathscr{R} \to N$ is called *limiting partial recursive* if and only if there is a partial recursive function ϕ such that for all $f \in \mathscr{R}$, $\psi(f) = \lim_{n \to \infty} \phi(f \mid n)$. We will call the function ϕ an *approximation* to ψ. If ϕ is a total recursive function then ψ will be a *limiting recursive functional* in the sense of Gold (1965), since n can be effectively obtained from (the encoding of) $f \mid n$. Clearly, every limiting recursive functional has some recursive approximation. A limiting recursive functional ψ is called *total* if $\psi(f)\downarrow$ for all $f \in \mathscr{R}$. We adopt the convention of indexing limiting partial recursive functionals according to their underlying partial recursive approximations, i.e., by ϕ_i: $\mathscr{R} \to N$ we mean the limiting partial recursive functional $\psi = \phi_i$ such that $\psi(f) = \lim_{n \to \infty} \phi_i(f \mid n)$. In this way we see that for any IIM M the (partial) map $\phi_M: \mathscr{R} \to N$ is a limiting partial recursive functional. For each limiting partial recursive functional ϕ_i and each $f \in \mathscr{R}$, we define the *modulus functional* μ_i (see Schoenfield, 1971) by,

$$\mu_i(f) = \begin{cases} m, & \text{where } m \text{ is the least integer such that} \\ & \phi_i(f \mid m)\downarrow \text{ and for all } n > m \text{ either } \phi_i(f \mid n)\uparrow \\ & \text{or } \phi_i(f \mid n)\downarrow = \phi_i(f \mid m), \\ \uparrow, & \text{if no such } m \text{ exists.} \end{cases}$$

Thus we see that $\mu_i(f)$ is the point of convergence of ϕ_i on f, that $\phi_i(f) = \phi_i(f \mid \mu_i(f))$, and that μ_i is a limiting recursive functional. Our results for limiting recursive functionals will not depend on the particular total recursive approximation ϕ_i chosen, so we will write $\psi(f \mid n)$ to mean $\phi_i(f \mid n)$, where ϕ_i is some a priori fixed total recursive approximation to ψ. We also write $\mu(\psi, f)$ for $\mu_i(f)$. Suppose $\psi = \phi_i$ is a limiting recursive functional, and define the partial recursive function ϕ_j by

$$\phi_j(\sigma) = \begin{cases} \phi_i(\sigma \mid (\max_\sigma/2)), & \text{if } \max_\sigma \text{ is even,} \\ \uparrow, & \text{if } \max_\sigma \text{ is odd.} \end{cases}$$

Then $\psi(f) = \lim_{n\to\infty} \phi_i(f \mid n) = \lim_{n\to\infty} \phi_j(f \mid n)$, so that ϕ_j is also an approximation to ψ. The fact that an IIM, as this example illustrates, can have many undefined responses to the initial segments of a function and still converge to a correct program for it will cause several of the definitions below to be somewhat complex. The reader should keep this particular situation in mind in going through the constructions below. To aid the readability of these definitions we define $D_{i,\sigma}$ to mean $\{n \leqslant \max_\sigma \mid \phi_i(\sigma \mid n)\downarrow\}$, and $D_{i,f}$ to mean $\{n \mid \phi_i(f \mid n)\downarrow\}$. We define the number of mind changes made by an IIM ϕ_M on input function f by

$$\delta_M(f) = \text{card}\{n \mid \phi_M(f \mid n)\downarrow \neq \phi_M(f \mid \max D_{M, f \mid n-1})\},$$

and if $\psi = \phi_i$ then we write $\delta(\psi, f)$ for $\delta_i(f)$. One can easily see that δ_M is a limiting recursive functional. Finally, the number of distinct hypotheses produced by an IIM M on input f is given by $\text{card}(\{\phi_M(f \mid n) \mid n \in D_{M,f}\})$.

To simplify subsequent arithmetic expressions assume that ϕ_0 is everywhere undefined, so there is no need to diagonalize against it. Also, $\phi_M(\dot{\phi})$ will denote an initial guess made by the IIM M before it examines any part of the input function. Most of the constructions below will involve functions of finite support (i.e., functions which are nonzero at only finitely many points). The set of functions of finite support will be denoted by $\mathscr{S}_*$ and defined formally by

$$\mathscr{S}_* = \{f \in \mathscr{R} \mid f =^* 0\}.$$

The functions of finite support provide a suitable class of functions for diagonalizing against IIMs, since each IIM which converges on an input function must do so on some finite initial segment of that function. Indeed, for many of the IIMs constructed below the programs hypothesized by them will take the simple form p_σ for some σ, where

$$\phi_{p_\sigma}(x) = \begin{cases} \sigma(x), & \text{if } x \in \operatorname{dom} \sigma, \\ 0, & \text{otherwise.} \end{cases}$$

In particular, there is a natural IIM M_* which infers the set $\mathscr{S}_*$ and which is defined by $\phi_{M_*}(\phi) = p_\phi$, and

$$\phi_{M_*}(\sigma) = \begin{cases} p_\sigma, & \text{if } \sigma(\max_\sigma) \neq 0, \\ \phi_{M_*}(\sigma - 1), & \text{otherwise.} \end{cases}$$

2. Complexity of Inference

For ordinary computations the complexity of a computation is synonymous with the complexity of the mechanism performing the computation. IIMs, however, are continually receiving inputs and reevaluating their most recent conjecture as to whether or not it is a program which computes the input function. Any IIM will then use an infinite amount of time to attempt any inference independently of whether or not the IIM eventually converges. The notion of inference complexity introduced below will measure only the amount of computation resources used up to the point of convergence. Initially a particular kind of complexity measure is introduced and some trade-offs between complexity and accuracy are exhibited.

Suppose $\{\tilde{\Phi}_i\}$ is any Blum computational complexity measure (Blum, 1967), for $\{\phi_i\}$ which is an acceptable numbering, so that

(1) $\tilde{\Phi}_i(x)\downarrow \Leftrightarrow \phi_i(x)\downarrow$

(2) $\tilde{\Phi}_i(x) = y$ is a recursive predicate in i, x, and y.

We extend this definition to limiting recursive functionals, and define the *inference complexity* of an IIM M on input f by

$$\Phi_M(f) = \sum_{n \in I} \tilde{\Phi}_M(f \mid n),$$

where $I = \{n \leqslant \mu_M(f) \mid n \in D_{M,f}\}$. Intuitively, since $\tilde{\Phi}_M(f \mid n)$ represents the effort expended by M in arriving at its response to $f \mid n$, the above measure represents the sum of the computation resources used by M on input f up to the point at which M converges to a particular program, and thus is the "area under the curve" of effort during the active period of inference by M.

Observe that only the computational complexity of M on inputs for which it conjectured a program is included in the summation. Clearly, this is a natural notion of inference complexity. Accordingly, we will call such an inference complexity measure an *a.u.c. measure*. Observe that so long as M converges the complexity of M is defined whether or not the program to which M converged is a correct program or f. This is analogous to the situation in analysis of algorithms where proofs of correctness and derivation of complexity bounds are treated separately. We will assume that $\tilde{\Phi}_M(\sigma) \geqslant \max_\sigma$ holds for the underlying computational complexity measure, so that $\Phi_M(f) \geqslant \mu_M(f)$.

Our first result shows the existence of a precise trade-off between complexity and accuracy. Particular subsets of the functions of finite support are shown to be inferrable either quickly or accurately, but not both. Viewed another way, for some classes of functions, increased accuracy can be obtained, but only with a proportionate increase in inference complexity.

THEOREM 1. *There exists a $g \in \mathscr{R}$ and IIMs $M_0, M_1, \ldots$, such that for all monotonically increasing $h \in \mathscr{R}$ and for all $n \in N$ there exists $S_{n,h} \in EX^n$ and*

(1) *for all $k < n$, $S_{n,h} \subseteq \mathscr{S}_*$ and $S_{n,h} \subseteq EX^{n-k}(M_k)$ and $(\forall^\infty f \in \mathscr{S}_{n,h})$ $[h(\Phi_{M_{k-1}}(f)) \leqslant \Phi_{M_k}(f) \leqslant g(h(\Phi_{M_{k-1}}(f)))]$,*

(2) *for all M and for all $k \leqslant n$ and for all $m < k$, if $\mathscr{S}_{n,h} \subseteq EX^m(M)$ then $(\forall f \in \mathscr{S}_{n,h})[\Phi_M(f) > h^{k-m}(\Phi_{M_m}(f))]$.*

Proof. Define M_k for $k \geqslant 0$ as follows:
$\phi_{M_k}(\varnothing) = 0$ and for σ extending $\varnothing$,

$$\phi_{M_k}(\sigma) = \begin{cases} p_\sigma, & \text{if } \sigma(\max_\sigma) \neq 0 \text{ and } \operatorname{card}\{x \in \operatorname{dom} \sigma \mid \sigma(x) \neq 0\} \leqslant k, \\ \phi_{M_k}(\sigma - 1), & \text{otherwise.} \end{cases}$$

Observe that if $f \in \mathscr{R}$, then $\phi_{M_k}(f \mid n)\downarrow$ for all n and $\phi_{M_k}(f)\downarrow$ and so $\Phi_{M_k}(f)\downarrow$. In essence, M_k hypothesizes that the input function has at most k nonzero values. Thus, for any f such that $f =^n 0$, where $n \geqslant k$, we will have $f \in EX^{n-k}(M_k)$. Define the function g by

$$g(z) = \max\left\{\sum_{i=0}^{z} \Phi_{M_k}(\sigma \mid i) \mid k \leqslant z \text{ and } \max_\sigma \leqslant z \text{ and } \operatorname{ran} \sigma \leqslant z^2 + 1\right\}.$$

Let $\mathscr{S}_{n,h} = \{f_{n,h,j} \mid j \geqslant 0\}$, where $f_{n,h,j}$ is the function of finite support non-effectively defined below. The function $f_{n,h,j}$ will have a nonzero value on arguments $x_{j,1}, \ldots, x_{j,n}$ which are defined iteratively. Since M_0 always outputs a singly conjecture, its complexity is constant. Also, M_{i-1} requires only $i-1$ nonzero points in the range of its input function to converge.

Consequently, the value $\Phi_{M_{i-1}}$ can be determined from $x_{j,1},\dots, x_{j,i-1}$. We now define

$$x_{j,i} = h(\Phi_{M_{i-1}}(f_{n,h,j})), \qquad \text{for} \quad 1 \leqslant i \leqslant n,$$

$$f_{n,h,j}(x) = \begin{cases} 0, & \text{if} \quad x \neq x_{j,i} \text{ for } 1 \leqslant i \leqslant n, \\ \min\{u \geqslant 1 \mid (\forall M \leqslant j)[\phi_{\phi_M(f_{n,h,j} \mid y_{j,i,M})}(x) \neq u]\}, & \\ \qquad \text{if} \quad x = x_{j,i} \text{ for some } 1 \leqslant i \leqslant n, & \end{cases}$$

where $y_{j,i,M} = \max\{y < x_{j,i} \mid \phi_M(f \mid y)\downarrow\}$.

By the definition of $f_{n,h,j}$ since $M \leqslant j$ and $i \leqslant n$ we have $\operatorname{card}\{\phi_{\phi_M(f_{n,h,j} \mid y_{j,i,M})}(x_{j,i}) \mid M \leqslant j \text{ and } i \leqslant n\} \leqslant n \times j + 1$, so that $f_{n,h,j}(x_{j,i}) \leqslant n \times j + 2$ for all $1 \leqslant i \leqslant n$, and $f_{n,h,j} =^n 0$ and hence $f_{n,h,j} \in \mathscr{S}_* \subseteq \mathscr{R}$.

Given M suppose j is so large that $j > \max\{M, \Phi_M(\phi), n\}$. If $\phi_M(f_{n,h,j})\downarrow$ and $\mu_M(f_{n,h,j}) < x_{j,k}$, then $\phi_M(f_{n,h,j} \mid y_{j,i,M}) = \phi_M(f_{n,h,j})$ for $k \leqslant i \leqslant n$, and so by the definition of $f_{n,h,j}$, $\phi_{\phi_M(f_{n,h,j})}(x_{j,i}) \neq f_{n,h,j}(x_{j,i})$ for $k \leqslant i \leqslant n$, and thus $f_{n,h,j} \notin EX^{n-k}(M)$. Thus, if $f_{n,h,j} \in EX^{n-k}(M)$ then $\mu_M(f_{n,h,j}) \geqslant x_{j,k}$, and so $\Phi_M(f_{n,h,j}) \geqslant x_{j,k}$. Since each $f_{n,h,j} =^n 0$, by our remarks above concerning M_k we see that $f_{n,h,j} \in EX^{n-k}(M_k)$. Thus, we have $\Phi_{M_k}(f_{n,h,j}) \geqslant x_{j,k} = h(\Phi_{M_{k-1}}(f_{n,h,j}))$ for all $1 \leqslant k \leqslant n$. Therefore, for any $m < k$, iterating this inequality we have $\Phi_M(f_{n,h,j}) > h^{k-m}(\Phi_{M_m}(f_{n,h,j}))$. Finally, it is clear that $\mu_{M_k}(f_{n,h,j}) = x_{j,k} = h(\Phi_{M_{k-1}}(f_{n,h,j}))$, and therefore since $\operatorname{ran} f_{n,h,j} \leqslant n \times j + 2$ and $x_{j,k} > j > n \geqslant k$, we have $\Phi_{M_k}(f_{n,h,j}) = \sum_{i=0}^{x_{j,k}} \Phi_{M_k}(f_{n,h,j} \mid i) \leqslant g(x_{j,k})$. ∎

The main reason for the complicated nature of the construction in Theorem 1 was the quantification $(\forall^\infty f \in \mathscr{S}_{n,h})$. Simpler sets of 0–1 valued functions can be constructed for the weaker case where the quantification $(\exists^\infty f \in \mathscr{S}_{n,h})$ is used. The results of the next section on the axiomatic approach indicate that there are some difficulties in establishing almost everywhere results for the complexity of inductive inference. This is especially true for results involving upper bounds on complexity, which differs dramatically from the situation regarding the computational complexity of the recursive functions.

Our next theorem reinforces our contention that such strong trade-offs as depicted in Theorem 1 are indeed rare. Let $\bar{\mathscr{S}}_{n,h} = \bigcup_{k=1}^{n} \mathscr{S}_{k,h}$. Since $\mathscr{S}_{k,h} \subseteq \mathscr{S}_*$, then clearly so is $\bar{\mathscr{S}}_{n,h}$.

THEOREM 2. *There exist IIMs $\bar{M}_1, \bar{M}_2,\dots,$ such that for all $n \in N$ there exists a constant c such that for all $k \in N$ $\bar{\mathscr{S}}_{n,h} \subseteq EX^k(\bar{M}_k)$ and $(\exists^\infty f \in \bar{\mathscr{S}}_{n,h})$ $[\Phi_{\bar{M}_k}(f) \leqslant c]$.*

Proof. We define $\bar{M}_k$ as follows:

$\phi_{\bar{M}_k}(\varnothing) = p_\varnothing$, and

$$\phi_{\bar{M}_k}(\sigma) = \begin{cases} p_\sigma, & \text{if } \sigma(\max_\sigma) \neq 0 \text{ and } \operatorname{card}\{x \in \operatorname{dom} \sigma \mid \sigma(x) \neq 0\} > k, \\ p_\varnothing, & \text{otherwise.} \end{cases}$$

Thus $\bar{M}_k$ is a modified version of M_* which ignores the first k nonzero values of the input function. Clearly, $\mathscr{S}_* \subseteq EX^k(\bar{M}_k)$ since $\bar{M}_k$ behaves like M_* after the first k nonzero values, so that $\bar{\mathscr{S}}_{n,h} \subseteq EX^k(\bar{M}_k)$ as well. Let $c_k = \Phi_{\bar{M}_k}(\varnothing)$ and $c = \max\{c_k \mid 1 \leqslant k \leqslant n\}$. Now, $\mathscr{S}_{j,h}$ is infinite for any $1 \leqslant j \leqslant k$, and for any $f \in \mathscr{S}_{j,h}$ we have that $\Phi_{\bar{M}_k}(\varnothing)$ is a k-variant of f and, because of the delay built into $\bar{M}_k$, that $\Phi_{\bar{M}_k}(f) \leqslant c_k$. Therefore, $(\exists^\infty f \in \bar{\mathscr{S}}_{n,h})[\Phi_{\bar{M}_k}(f) \leqslant c]$. ∎

Theorem 2 demonstrates the existence of easy to infer sets of functions. We conclude this section with a result which emphasizes this point.

THEOREM 3. *There exists a set of total functions $\mathscr{S}$ and an IIM M and a total recursive function h such that $\mathscr{S} \subseteq EX(M)$ and $(\forall f \in \mathscr{S})[\Phi_M(f) \leqslant h(f(0))]$ and $(\forall f \in \mathscr{R})(\exists g \in \mathscr{S})[f =^1 g]$.*

Proof. Let $\mathscr{S} = \{f \mid \phi_{f(0)} = f\}$, and $\phi_M(\sigma) = \sigma(0)$, and $h(z) = \Phi_M(\langle(0, z)\rangle)$. Then clearly $\mathscr{S} \subseteq EX(M)$ and $\Phi_M(f) \leqslant \Phi_M(f \mid 0) = h(f(0))$ for all $f \in \mathscr{S}$. Also, if $\phi_e \in \mathscr{R}$, then via the recursion theorem we can define

$$\phi_{\alpha(e)}(x) = \begin{cases} \alpha(e), & \text{if } x = 0, \\ \phi_e(x), & \text{otherwise.} \end{cases}$$

Clearly, $\phi_{\alpha(e)} \in \mathscr{S}$ and $\phi_{\alpha(e)}(x) = \phi_e(x)$ for all $x > 0$, so that $\phi_{\alpha(e)} =^1 \phi_e$. ∎

3. Axiomatic Approach

In (Blum, 1967) he initiates an axiomatic approach to the the complexity of computing partial recursive functions. His approach has proved to be a tremendous success in understanding the nature of computations and their complexity. In this section we formulate an analogous axiomatization of the complexity of inductive inference. In some sense the a.u.c. measures of the previous section could be considered an axiomatization since they were based on an arbitrary Blum computational complexity measure. However, to restrict attention only to such measures would exclude from consideration the modulus function and the number of mind changes as possible measures of inference complexity. Since a priori these seem to represent reasonable notions of measure, we will attempt an axiomatization which includes them. It may be the case, though, that just as the axiomatization of Blum (necessarily) admitted pathological measures, our axiomatization may also allow some pathologies. Consequently, it may be useful at some later point to restrict consideration to more natural measures such as the a.u.c. measures.

Freivalds (1975) noted a natural and straightforward way to extend the notion of Blum computational complexity measures to the limiting recursive functions. He combined the well-known result that limiting recursive functions are equivalent to functions computable using the Halting Problem as an oracle with the work of Khodzhayev (1970) (see also Lynch, Meyer, and Fischer, 1976; Symes, 1971), who extended the notion of Blum computational complexity to oracle computations and showed that many results from the complexity theory of ordinary computations (e.g., the existence of arbitrarily difficult to compute functions, the speed-up theorem) also hold for oracle computations. Freivalds also observed that one of the commonly used measures of complexity of inference, viz. the number of distinct hypotheses, when applied to limiting recursive functions failed to satisfy the first Blum axiom (since there can be an infinite alternation among a finite number of values). One, of course, could prevent this by restricting attention to only those IIMs which once they have rejected a hypothesis (i.e., have changed their mind from it to another hypothesis) never return to it. Freivalds then presented a characterization of the type of complexity measures suitable for the number of distinct hypotheses. Shäfer-Richter (1984) has presented recently an axiomatization of the complexity of inductive inference which is different from that given below in two respects. First, the measure deals with presentations of recursive functions instead of the functions themselves, and second, the analog of Blum's second axiom states essentially that the set of triples (i, σ, y) such that $\Phi_i(\sigma) \geqslant y$ is recursively enumerable. Our axiomatization will also be a natural extension of the axioms of Blum but to limiting recursive functionals. The fact that a functional has a function space as its domain (which does not have the nice order properties that the integers do) will shape many of the results which follow, some of which contrast with those obtainable for the complexity of ordinary computations.

With these introductory remarks in mind we now present our axiomatization of an inference complexity measure. We say that a set of functionals $\{\Phi_i\}$ is an inference complexity measure for the acceptable Gödel numbering $\{\phi_i\}$ if and only if the following are axioms satisfied:

AXIOM 1. $\Phi_i(f)\downarrow \Leftrightarrow \phi_i(f)\downarrow$.

AXIOM 2. There exists a total limiting recursive functional Γ: $N \times \mathscr{R} \times N \to N$ such that

$$\Gamma(i, f, y) = \begin{cases} 1, & \text{if } \Phi_i(f) = y, \\ 0, & \text{otherwise.} \end{cases}$$

Note that we do not require Φ_i to be a limiting partial recursive functional, although we will soon see that this follows from Axiom 2. Also, it is possible that $\Phi_M(f)\downarrow$ and yet $f \notin EX(M)$, so that the complexity of inference will be distinct from the correctness of an inference. Axioms 1 and 2 represent a natural extension of the Blum computational complexity axioms to limiting recursive functionals and are in essence the same as those given by Freivalds (1975) for limiting recursive functions.

The following lemma is an analog of a similar result for computational complexity measures and is fundamental to our notion of inference complexity.

LEMMA 4. *Φ_i is a limiting recursive functional.*

Proof. Let Γ be the total limiting recursive predicate given by Axiom 2. Then there is a total recursive approximation γ to Γ such that $\Gamma(i, f, y) = \lim_{n \to \infty} \gamma(i, f \mid n, y)$. Given Φ_i we define γ_i by

$$\gamma_i(\sigma) = \begin{cases} \min\{ y \leqslant \max_\sigma \mid \gamma(i, \sigma, y) = 1 \}, \\ \max_\sigma, \qquad \text{if no such } y \text{ exists.} \end{cases}$$

Clearly, γ_i is a total recursive function. Suppose $\Phi_i(f)\downarrow$, and let $y_0 = \Phi_i(f)$. Then, $\lim_{n \to \infty} \gamma(i, f \mid n, y_0) = 1$ and for all $y \neq y_0$, $\lim_{n \to \infty} \gamma(i, f \mid n, y) = 0$. Thus, $\lim_{n \to \infty} \gamma_i(f \mid n) = \lim_{n \to \infty} \min\{ y \leqslant n \mid \gamma(i, f \mid n, y) = 1 \} = y_0$, and therefore $\Phi_i(f) = \lim_{n \to \infty} \gamma_i(f \mid n)$. On the other hand, if $\Phi_i(f)\uparrow$ then $\Phi_i(f) \neq y$ for all y, and $\lim_{n \to \infty} \gamma(i, f \mid n, y) = 0$ for all y. Thus, $\lim_{n \to \infty} \gamma_i(f \mid n) = \infty$. Therefore, γ_i is a total recursive approximation to Φ_i, and hence Φ_i is a limiting recursive functional. ∎

It is clear from the proof of Lemma 4 that γ provides a uniform way of approximating any Φ_i, analagous to the situation for computational complexity measures, and we will henceforth use γ_i as the canonical approximation for Φ_i and write $\Phi_i(f \mid n)$ for $\gamma_i(f \mid n)$, $\mu(\Phi_i, f)$ for $\mu(\gamma_i, f)$, and $\delta(\Phi_i, f)$ for $\delta(\gamma_i, f)$. It is clear that any a.u.c. measure, the modulus functional μ_M and the number of mind changes functional δ_M all satisfy Axiom 1 above. To see that these measures satisfy Axiom 2 we first define an approximation to $D_{i,\sigma}$ by $\tilde{D}_{i,\sigma} = \{ n \leqslant \max_\sigma \mid \tilde{\Phi}_i(\sigma \mid n) \leqslant \max_\sigma \}$, where $\tilde{\Phi}_i$ is some a priori fixed computational complexity measure. In the specific case of the modulus functional we next define the total recursive approximation γ^{μ} to Γ^{μ} by,

$$\gamma^{\mu}(i, \sigma, y) = \begin{cases} 1, & \text{if } y \text{ is the least integer } m \in \tilde{D}_{i,\sigma} \text{ such that for all} \\ & n \in \tilde{D}_{i,\sigma} \text{ if } n > m \text{ then } \phi_i(\sigma \mid n) = \sigma_i(\sigma \mid m), \\ 0, & \text{if no such } m \text{ exists.} \end{cases}$$

The verification that γ^{μ} is a recursive approximation to Γ^{μ} is straightforward. In keeping with our notational convention we will write μ_i for γ^{μ_i}. Suppose next that Φ_i is the a.u.c. measure with underlying computational complexity measure $\tilde{\Phi}_i$ then it is not difficult to see that the function $\gamma^{\text{a.u.c.}}$ defined by

$$\gamma^{\text{a.u.c.}}(i, \sigma, y) = \begin{cases} 1, & \text{if } \; y = \sum_{n \in I} \tilde{\Phi}_M(\sigma \mid n), \text{ where} \\ & I = \{n \leqslant \mu_i(\sigma) \mid n \in \tilde{D}_{i,\sigma}\}, \\ 0, & \text{otherwise,} \end{cases}$$

is a recursive approximation to $\Gamma^{\text{a.u.c.}}$. One can analagously define a total recursive function γ^{δ}.

Axiom 2 is actually quite a powerful constraint on possible inference complexity measures. For example, if we define

$$\Phi_i(f) = \begin{cases} 0, & \text{if } \; \delta_i(f) \text{ is even,} \\ 1, & \text{if } \; \delta_i(f) \text{ is odd,} \\ \uparrow, & \text{if } \; \delta_i(f)\uparrow, \end{cases}$$

then clearly Φ_i is a limiting recursive functional and satisfies Axiom 1. We now show that it cannot satisfy Axiom 2. Suppose there is a total recursive approximation γ, as provided by Axiom 2, such that

$$\lim_{n \to \infty} \gamma(i, f \mid n, y) = \begin{cases} 1, & \text{if } \; \Phi_i(f) = y, \\ 0, & \text{if } \; \Phi_i(f) \neq y. \end{cases}$$

Consider the IIM M_* which identifies $\mathscr{S}_*$. We define a sequence $\{f_n\}$ of functions of finite support such that $\text{card}\{x \mid f_n(x) \neq 0\} = n$ and $\{x \mid f_n(x) \neq 0\} \subseteq \{x \mid f_{n+1}(x) \neq 0\}$ as follows: define $f_0(x) = 0$, for all x, and

x_0 to be the least integer x such that $\gamma(M_*, f_0 \mid x, 0) = 1$ and $\gamma(M_*, f_0 \mid x, 1) = 0$. Then given f_n and x_n we define

$$f_{n+1}(x) = \begin{cases} f_n(x), & \text{if } x \leqslant x_n, \\ 1, & \text{if } x = x_n + 1, \\ 0, & \text{if } x > x_n + 1, \end{cases}$$

and x_{n+1} to be the least integer x such that $\gamma(M_*, f_n \mid x, e) = 1$ and $\gamma(M_* f_n \mid x, 1 - e) = 0$, where $e = n \bmod 2$. Finally, define the function $f_\infty \in \mathscr{R}$ by

$$f_\infty(x) = \begin{cases} 1, & \text{if } f_n(x) = 1 \text{ for some } n, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, $f_\infty \notin \mathscr{S}_*$ and for all n, $f_\infty \mid x_n = f_n \mid x_n$. Then for all even n we have $\gamma(M_*, f_\infty \mid x_n, 0) = \gamma(M_* f_n \mid x_n, 0) = 1$ and $\gamma(M_*, f_\infty \mid x_n, 1) = \gamma(M_*, f_n \mid x_n, 1) = 0$, and for odd n we have $\gamma(M_*, f_\infty \mid x_n, 1) = 1$ and $\gamma(M_*, f_\infty \mid x_n, 0) = 0$. Therefore, neither $\lim_{n \to \infty} \gamma(M_* f_\infty \mid n, 0)$ nor $\lim_{n \to \infty} \gamma(M_* f_\infty \mid n, 1)$ exist, so that no such total limiting recursive functional Γ can exist. That the above should not be an inference complexity measure can be seen from the fact that if ϕ_i is a total recursive function then $\Phi_i(f) \leqslant 1$ for all $f \in \mathscr{R}$, so that the predicate $P_i(f) \equiv \Phi_i(f) \leqslant 1$ is total limiting recursive and identically true. One is reminded here of the difficulty in axiomatic computational complexity where a program can cycle indefinitely using only a finite amount of memory space. We observe further that if $\phi_i(f)\uparrow$ then $\mu_i(f) = \infty$ (and $\delta_i(f) = \infty$ as well), and so $\Phi_i(f) = \infty$. This observation is crucial in several of the constructions below where this fact that an IIM cannot change its mind indefinitely without increasing its complexity of inference is used. It is also easy to see by this example that the number of distinct hypotheses produced by an IIM cannot serve as an inference complexity measure.

Even though the modulus functional is an inference complexity measure, the modulus functional cannot be an a.u.c. measure (see the discussion below following Lemma 5 for details). Thus not every inference complexity measure is an a.u.c. measure.

One of the very rich areas of work in computational complexity was in the study of complexity classes. We analagously define for any limiting (partial) recursive function ψ the inference complexity class named by ψ by

$$C_\psi = \{\mathscr{S} \mid (\exists M)\,(\forall^\infty f \in \mathscr{S})\,[\Phi_M(f) \leqslant \psi(f)]\}.$$

Note that if $f \notin EX(M)$, then $\Phi_M(f)$ is unbounded. Only total limiting recursive functionals are used for names of complexity classes below. One of the most useful and pleasing properties of computational complexity measures is their recursive relatedness, viz. if $\{\Phi_i\}$ and $\{\hat{\Phi}_i\}$ are two computational complexity measures then there exists a total recursive function r such that for all $n \in N$ and for all $i \geqslant n$, $\Phi_i(n) \leqslant r(n, \hat{\Phi}_i(n))$ and $\hat{\Phi}_i(n) \leqslant r(n, \Phi_i(n))$. Below is our analog of recursive relatedness for inference complexity measures.

LEMMA 5. *If $\{\Phi_i\}$ and $\{\hat{\Phi}_i\}$ are two inference complexity measures for the acceptable Gödel numbering $\{\phi_i\}$, then there exists a total limiting recursive functional Ψ: $\mathscr{R} \times N \times N \to \mathscr{R}$ such that for all $i \in N$ and for all $f \in \mathscr{R}$ $\Phi_i(f) \leqslant \Psi(f, \hat{\Phi}_i(f), i)$ and $\hat{\Phi}_i(f) \leqslant \Psi(f, \Phi_i(f), i)$.*

Proof. Define

$$\Psi(\sigma, s, i) = \max\{\Phi_i(\sigma), \hat{\Phi}_i(\sigma) \mid \text{either } \Phi_i(\sigma) \leqslant s \text{ or } \hat{\Phi}_i(\sigma) \leqslant s\}.$$

Fix s, i, and f. Suppose $\Phi_i(f)\uparrow$. Then by Axiom 1, $\phi_i(f)\uparrow$ and $\hat{\Phi}_i(f)\uparrow$ also. Then $\lim_{n\to\infty} \Phi_i(f \mid n) = \lim_{n\to\infty} \hat{\Phi}_i(f \mid n) = \infty$ so that for all but finitely many integers n $\Phi_i(f \mid n) \neq s$ and $\hat{\Phi}_i(f \mid n) \neq s$, and hence $\Psi(f \mid n, s, i) = \max \phi = 0$. Therefore, $\Psi(f, s, i) = 0$. Suppose now that $\Phi_i(f)\downarrow$. Then again by Axiom 1 $\phi_i(f)\downarrow$ and $\hat{\Phi}_i(f)\downarrow$. If $\Phi_i(f) \neq s$ and $\hat{\phi}_i(f) \neq s$, then for all but finitely many integers n, $\Psi(f \mid n, s, i) = \max \phi = 0$ so that as before $\Psi(f, s, i) = 0$. If either $\Phi_i(f) = s$ or $\hat{\Phi}_i(f) = s$, then for all but finitely many integers n, $\Psi(f \mid n, s, i) = \max\{\Phi_i(f \mid n), \hat{\Phi}_i(f \mid n)\}$ so that $\lim_{n\to\infty} \Psi(f \mid n, s, i)$ exists and $\Psi(f, s, i) = \max\{\Phi_i(f), \hat{\Phi}_i(f)\}$. In this case clearly $\Psi(f, \Phi_i(f), i) \geqslant \hat{\Phi}_i(f)$ and $\Psi(f, \hat{\Phi}_i(f), i) \geqslant \Phi_i(f)$. In all cases $\Psi(f, s, i)$ is defined so that Ψ is a total limiting recursive functional. ∎

Observe that the recursive relatedness of Φ_i and $\hat{\Phi}_i$ depends on i in contrast to the case for computational complexity measures. In an attempt to obtain the kind of almost everywhere result seen for computational complexity measures, viz., where there exists a total limiting recursive functional $\tilde{\Psi}$: $\mathscr{R} \times N \to N$ satisfying $\Phi_i(f) \leqslant \tilde{\Psi}(f, \hat{\Phi}_i(f))$ and $\hat{\Phi}_i(f) \leqslant \hat{\Psi}(f, \Phi_i(f))$, one might try to define such a total limiting recursive functional $\tilde{\Psi}$ by $\tilde{\Psi}(\sigma, s) = \max\{\Psi(\sigma, s, i) \mid i \leqslant \max_\sigma\}$. However, the following example shows that this is not possible. Let $\{\Phi_i\}$ be an a.u.c. type inference complexity measure, and let $\hat{\Phi}_i(f) = \mu_i(f)$. Let $i_1, i_2, i_3,\ldots,$ be a sequence of programs such that $\phi_{i_1} = \phi_{i_2} = \phi_{i_3} = ,\ldots,$ and $\Phi_{i_1} < \Phi_{i_2} < \Phi_{i_3} < ,\ldots.$ Such a sequence is easily constructed using well-known runtime padding techniques from computational complexity. Since $\mu_{i_1}(f) = \mu_{i_2}(f) = \mu_{i_3}(f) = ,\ldots,$ we see that $\hat{\Phi}_{i_1} = \hat{\Phi}_{i_2} = \hat{\Phi}_{i_3} = ,\ldots.$ Thus, if $\tilde{\Psi}$ is any limiting recursive functional such that $\Phi_i(f) \leqslant \tilde{\Psi}(f, \tilde{\Phi}_i(f))$ for all i and $f \in \mathscr{R}$, then $\tilde{\Psi}(f, \mu_{i_1}(f)) > \Phi_{i_k}(f)$ for all values i_k, so that $\tilde{\Psi}(f, u_{i_1}(f))\uparrow$ and $\tilde{\Psi}$ cannot be total. This contradiction also shows that $\{\mu_i\}$ cannot be an a.u.c. inference complexity measure.

An alternative formulation of $\tilde{\Psi}$ might be $\tilde{\Psi}(\sigma, s) = \max\{\Psi(\sigma, s, i) \mid i \leqslant s\}$. In this case we would have for all $i \in N$ and for all $f \in \mathscr{R}$, if $\mu_i(f) > i$ then $\Phi_i(f) \leqslant \Phi_i(f))$ and $\hat{\Phi}_i(f) \leqslant \hat{\Psi}(f, \Phi_i(f))$. But this is still not an almost everywhere result, since one can easily construct examples, where $\Phi_i(f) \leqslant i$ (and so $\mu_i(f) \leqslant i$) for infinitely many $f \in \mathscr{R}$. Moreover, we have

LEMMA 6. *There does not exist a total limiting recursive functional ψ such that* $(\forall n \in N)(\forall^\infty f \in \mathscr{R})\ [\psi(f) > n]$.

Proof. Suppose ψ is a total limiting recursive functional. We construct an infinite set of programs $\{p_{n,j}\}$ for distinct total recursive functions. The construction proceeds in stages and at stage n all the programs $p_{n,j}$ will be defined and will satisfy $\forall^\infty x[\phi_{p_{n,j}}(x) = j]$. Let $\langle \cdot, \cdot \rangle$ be an effective enumeration of all pairs of integers, and define $\sigma_0 = \varnothing$ and $v_0 = 0$.

Stage n. Define

$$\phi_{p_{n,j}}(x) = \begin{cases} \sigma_n(x), & \text{if} \quad x \in \operatorname{dom} \sigma_n, \\ j, & \text{otherwise.} \end{cases}$$

Look for the least pair $\langle j, x \rangle$ such that $x \notin \operatorname{dom} \sigma_n$ and $\psi(\phi_{p_{n,j}} \mid x) > v_n$. When and if such j and x are found set $\sigma_{n+1} = \phi_{p_{n,j}} \mid x$ and $v_{n+1} = \psi(\phi_{p_{n,j}} \mid x)$ and then go to stage $n+1$. End stage n. ▮

Each program $p_{n,j}$ computes a finite variant of a constant function hence each $\phi_{p_{n,j}} \in \mathscr{R}$. The effectiveness of the construction lies in the fact that $p_{n,j}$ depends only on σ_n and j. Either the search for x and j succeeds and only finitely many programs are generated or the search fails and the construction snags at stage n generating and testing programs $p_{n,0}, p_{n,1},\ldots$ If there are finitely many stages in the construction, then since $\sigma_n \subset \sigma_{n+1}$ the function $f = \bigcup_{n=0}^\infty \sigma_n$ will be total recursive, and since $\psi(\sigma_{n+1}) > \psi(\sigma_n)$ for all n, we will have $\psi(f)\uparrow$, so ψ cannot be total. If, on the other hand, there are only finitely many stages and stage n is the last, then for each $j \in N$ and each $x \notin \operatorname{dom} \sigma_n$ we have $\psi(\phi_{p_{n,j}} \mid x) \leqslant v_n$, and since ψ is total, $\psi(\phi_{p_{n,j}}) \leqslant v_n$. Thus, there exist infinitely many $f \in \mathscr{R}$ such that $\psi(f) \leqslant v_n$. ▮

Observe that if such a limiting recursive functional were to exist then several of the results from computational complexity theory would carry over to inference complexity (e.g., the Gap Theorem and the Honesty Theorem). The following result shows that even though the modulus measure $\{\mu_i\}$ cannot be an a.u.c. measure, for IIM's which are defined on all data inputs every a.u.c. measure can be levelled into the modulus measure, so that in some sense the modulus measure is a canonical measure.

LEMMA 7. *For every a.u.c. measure* $\{\Phi_i\}$ *there exists program transformation* $\alpha \in \mathcal{R}$ *such that for all i for which* ϕ_i *is a limiting recursive functional and for all* $f \in \mathcal{R}$, $\phi_{\alpha(i)}(f) = \phi_i(f)$ *and* $\delta_{\alpha(i)}(f) = \delta_i(f)$ and $\mu_{\alpha(i)}(f) = \Phi_i(f)$.

Proof. Let $\{\tilde{\Phi}_i\}$ be the underlying computational complexity measure upon which $\{\Phi_i\}$ is based. Define α as follows:

$$\phi_{\alpha(i)}(\sigma) = \phi_i(\sigma'),$$

where σ' is the largest initial segment of σ such that

$$\sum_{k=0}^{\max_{\sigma'}} \tilde{\Phi}_i(\sigma \mid k) \leqslant \max_{\sigma} \text{ and } \phi_i(\sigma')\downarrow \neq \phi_i(\sigma' - 1).$$

If $\phi_i(f)\uparrow$, then $\phi_{\alpha(i)}(f)\uparrow$ and $\Phi_i(f) = \mu_{\alpha(i)}(f) = \infty$. If $\phi_i(f)\downarrow$, then $\phi_{\alpha(i)}(f \mid n) = \phi_i(f \mid \mu_i(f)) = \phi_i(f)$ for all $n \geqslant \Phi_i(f)$, so that $\phi_{\alpha(i)}(f) = \phi_i(f)$. Also, if $n < \Phi_i(f)$, then $\phi_{\alpha(i)}(f \mid n) = \phi_i(f \mid m)$ for some $m < \mu_{\alpha(i)}(f)$, so that there will be some $k > n$ such that $\phi_{\alpha(i)}(f \mid k) \neq \phi_i(f \mid m)$ (recall here that we are assuming that $\tilde{\Phi}_i(x) > 0$), and thus $\mu_{\alpha(i)}(f) > n$. Therefore, $\Phi_i(f) = \mu_{\alpha(i)}(f)$. It is also clear that $\delta_{\alpha(i)}(f) = \delta_i(f)$, since $\phi_{\alpha(i)}$ outputs hypotheses in the same order (but with perhaps many repetitions) as ϕ_i. ∎

We now show that there are arbitrarily difficult to infer sets of total recursive functions.

THEOREM 8. *For every inference complexity measure* $\{\Phi_i\}$ *and for every total limiting recursive functional* ψ *there exists a set* $\mathcal{S}_\psi \subseteq \mathcal{S}_*$ *of total recursive functions such that* $\mathcal{S}_\psi \in EX$ *and* $(\forall M)$ $(\forall^\infty f \in \mathcal{S}_\psi)$ [*if* $f \in EX(M)$ *then* $\Phi_M(f) > \psi(f)$].

Proof. Let ψ be a total limiting recursive functional. We construct a set $\mathcal{S}_\psi = \{f_n\}$ of functions of finite support (so that $\mathcal{S}_\psi \subseteq EX(M_*)$) such that for all M and for all $n \geqslant M$ if $f_n \in EX(M)$ then $\Phi_M(f_n) > \psi(f_n)$. The function f_n is defined to be the 0-completion of the partial recursive function with cofinite domain $\phi_{\alpha(n)}$, i.e., if $\phi_{\alpha(n)}(x)\uparrow$ then $f_n(x) = 0$. The function $\phi_{\alpha(n)}$ is constructed in effective stages of finite extension in a manner similar to that given in Case and Smith (1983) to show that $EX \subset EX^1$. The construction of $\phi_{\alpha(n)}$ employs a movable marker whose position at stage s is given by a^s. Initially, $a^0 = 1$. We use σ^s to denote the largest initial segment of $\phi^s_{\alpha(n)}$ (the stage s approximation to $\phi_{\alpha(n)}$, i.e., the finite initial segment of $\phi_{\alpha(n)}$ defined prior to stage s) and x^s to denote the first integer $x > a^s$ such that $x \notin \operatorname{dom} \phi^s_{\alpha(n)}$. We fix some computational complexity measure $\{\tilde{\Phi}_i\}$. Also, for each $M \leqslant n$, we use p^s_M to denote M's most recently discovered conjecture for $\phi_{\alpha(n)}$ prior to stage s, i.e., $p^s_M = \phi_M(\sigma)$, where σ is the largest initial segment of $\phi^s_{\alpha(n)} \cup \{(a^s, 0)\}$ such that $\tilde{\Phi}_M(\sigma) \leqslant s$. Each $M \leqslant n$ is *marked* with one of **a**, **b**, or **c**. The marking **a** indicates that M's conjecture on the largest known initial segment of f_n has been rendered incorrect by a suitable extension to $\phi_{\alpha(n)}$ at some prior stage. When the inference complexity of M on the known initial segment of f_n exceeds the current estimate of $\psi(f_n)$, M's marking is changed to **b**. The initial marking **c** is used in all other situations indicating that M appears to be identifying f_n with too small a complexity. We initialize $\phi_{\alpha(n)}$ by setting $\phi_{\alpha(n)}(0) = n$, so that $\phi^0_{\alpha(n)} = \{(0, n)\}$. Since by convention ϕ_0 is everywhere undefined, there is no need to diagonalize against it. This initialization actually is not required in this theorem but rather in the analog of Blum's compression theorem which follows. We remark in passing that this same initialization technique was used by Barzdin (1974). We also remark that $\phi_{\alpha(n)}$ can be made to be 0–1 valued by encoding n as an initial string of n 1's in the graph of $\phi_{\alpha(n)}$ instead of setting $\phi_{\alpha(n)}(0) = n$.

Stage s (of $\phi_{\alpha(n)}$). Do the following three steps in order.

(1) If $\psi(\sigma^s) \neq \psi(\phi^s_{\alpha(n)} \cup \{(a^s, 0)\})$ or $\psi(\sigma^s) \neq \psi(\sigma^{s-1})$, then

(a) for all $M \leqslant n$ mark M **c**, and

(b) define $a^{s+1} = x^s$, and

(c) define $\phi^{s+1}_{\alpha(n)} = \phi^s_{\alpha(n)} \cup \{(a^s, 0)\}$, and

(d) go to stage $s+1$.

(2) For each $M \leqslant n$,

(a) if M is marked **b** and $\Phi_M(\sigma^s) \leqslant \psi(\sigma^s)$, then mark M **c**.

(b) if M is marked **c** and $\Phi_M(\sigma^s) > \psi(\sigma^s)$, then mark M **b**.

(3) (a) If for some $M \leqslant n$, M is marked **c** and $\tilde{\Phi}_{p^s_M}(a^s) \leqslant s$, then for the least such M

(i) mark M **a**, and

(ii) define $a^{s+1} = x^s$, and

(iii) define $\phi^{s+1}_{\alpha(n)} = \phi^s_{\alpha(n)} \cup \{(a^s, 1 \dot{-} \phi_{p^s_M}(a^s))\}$,

(b) else if for some $M \leqslant n$, M is marked **a** and $p^s_M \neq p^{s-1}_M$, then for the least such M

(i) mark M **c**, and

(ii) define $a^{s+1} = x^s$, and

(iii) define $\phi^{s+1}_{\alpha(n)} = \phi^s_{\alpha(n)} \cup \{(a^s, 0)\}$.

(c) else

(i) define $a^{s+1} = a^s$, and

(ii) define $\phi^{s+1}_{\alpha(n)} = \phi^{s}_{\alpha(n)} \cup \{(x^s, 0)\}$.

Go to stage $s+1$. *End stage s.*

We let f_n be the 0-completion of $\phi_{\alpha(n)}$, i.e., $f_n = \phi_{\alpha(n)} \cup \{(x, 0) \mid \phi_{\alpha(n)}(x)\uparrow\}$. Observe first that $\phi_{\alpha(n)}(x)\uparrow \Leftrightarrow x = \lim_{s\to\infty} a^s$, so that f_n will be total recursive. We first show that for each $M \leqslant n$, M is eventually permanently marked **a** or **b** or **c**. Observe that M cannot change its marking from **a** to **b** or from **b** to **a** directly, but must first be marked **c**. Observe also that in order for M to change its marking from **a** to **c** after stage $\mu(\psi, f)$, M must change its mind (see step 3b). If there are infinitely many **b**'s in M's marking sequence, then $\Phi_M(f_n \mid m) > \psi(f)$ for infinitely many m. If M is not permanently marked **b**, then $\Phi_M(f_n \mid m) \leqslant \psi(f)$ for infinitely many m also, but this contradicts the fact that if $\Phi_M(f_n)\uparrow$, then $\lim_{m\to\infty} \Phi_M(f \mid m) = \infty$. Suppose there are infinitely many **a**'s (and only finitely many **b**'s) in M's marking sequence and that M is not permanently marked **a**. Then M must be marked **c** infinitely many times, so that M must change its mind infinitely many times, and hence $\delta_M(f_n) = \infty$. But then $\Phi_M(f_n) = \infty$ also, so that M would be marked **b** at step 2b infinitely many times, which again is a contradiction. Therefore, every M will be permanently marked either **a** or **b** or **c**. Observe that in any event by case 3c, $a = \lim_{s\to\infty} a^s$ exists. If M is permanently marked **b**, then step 2a can only apply a finite number of times so that $\Phi_M(f_n \mid m) > \psi(f_n \mid m)$ for all but finitely many m, and $\Phi_M(f_n) > \psi(f_n)$. If M is permanently marked **a** or **c**, then M must stop changing its mind so that $\phi_M(f_n)\downarrow$. If M is permanently marked **a**, then eventually M's last program will be diagonalized against at step 3a and $\phi_{\phi_M(f_n)}(a) \neq f_n(a)$. If M is permanently marked **c**, then M's last program could not have been diagonalized at step 3a so that $\phi_{\phi_M(f_n)}(a)\uparrow$. Therefore, for each $M \leqslant n$ either $\phi_{\phi_M(f_n)} \neq f_n$ so $f_n \notin EX(M)$, or $\Phi_M(f_n) > \psi(f_n)$.

The referee has observed that the hypothesis in Theorem 8 (as well as Theorem 9 and 10 below) that ψ be a total limiting recursive functional can be weakened to the hypothesis that $\mathscr{S}_* \subseteq \operatorname{dom} \psi$. We now demonstrate an analog of Blum's compression theorem.

THEOREM 9. *For every inference complexity measure $\{\Phi_i\}$ there exists a total limiting recursive functional Ψ: $\mathscr{R} \times N \to \mathscr{R}$ such that for every total limiting recursive functional ϕ_i there exists a set $\mathscr{S}_{\phi_i} \subseteq \mathscr{S}_*$ of total recursive functions such that $\mathscr{S}_{\phi_i} \in EX$ and $\mathscr{S}_{\phi_i} \notin C_{\phi_i}$ and $\mathscr{S}_{\phi_i} \in C_\Delta$,* where $\Delta(f) = \Psi(f, \max\{\phi_i(f), \Phi_i(f)\})$.

Proof. Given a total limiting recursive functional ϕ_i we let $\mathscr{S}_{\phi_i}$ be the set of functions of finite support constructed in Theorem 8 above, where ϕ_i plays the role of ψ in the construction. Clearly, $\mathscr{S}_{\phi_i} \notin C_{\phi_i}$, and since $\mathscr{S}_{\phi_i} \subseteq \mathscr{S}_*$, $\mathscr{S}_{\phi_i} \subseteq EX(M_*)$. Because of the limiting recursive relatedness of inference complexity measures it suffices to construct a total limiting recursive functional Ψ such that $\delta_{M_*}(f) \leqslant \Psi(f, \max\{\phi_i(f), \mu_i(f)\})$ for all $f \in \mathscr{S}_{\phi_i}$. We define Ψ as follows:

$$\Psi(\sigma, s) = \sum_{M \leqslant \sigma(0)} \max\{s, 2 \times \delta_M(\sigma') + 1 \mid \sigma' \subseteq \sigma \text{ and } \Phi_M(\sigma') \leqslant s\}.$$

Clearly, Ψ is a total recursive function. To see that Ψ approximates a total limiting recursive functional fix f and s and let $M \leqslant f(0)$. If $\phi_M(f)\downarrow$, then $\delta_M(f)\downarrow$ and M's contribution to the sum defining $\Psi(f, s)$ is bounded by $2 \times \delta_M(f) + 1$. On the other hand, if $\phi_M(f)\uparrow$, then $\lim_{m \to \infty} \Phi_M(f \mid m) = \infty$ so only finitely many m satisfy $\Phi_M(f \mid m) \leqslant s$, and again M's contribution is bounded. Therefore, $\Psi(f, s)$ exists for all f and s. Consider $f_n \in \mathscr{S}_{\phi_i}$. Clearly, $\delta_{M_*}(f_n) = \text{card}\{x \mid f_n(x) \neq 0\}$. From the construction above for f_n it is clear that $f_n(x) \neq 0$ only if $x = a^s$ for some $s \geqslant 0$, and hence $\delta_{M_*}(f_n) \leqslant \text{card}\{a^s \mid s \geqslant 0\}$. For stages $s > \mu_i(f_n)$ the marker can move at stage s only because of step 3a or step 3b. Now, each $M \leqslant n$ can be responsible for at most $2 \times \delta_M(f) + 1$ marker movements, since each movement except the last requires M to change its marking from **a** to **c** at step 3b, and hence to change its mind, and then later to change its marking from **c** to **a** at step 3a. Therefore,

$$\begin{aligned}\delta_{M_*}(f_n) &\leqslant \sum_{M \leqslant n} \max\{\mu_i(f), 2 \times \delta_M(\sigma) + 1 \mid \sigma \subseteq f_n \text{ and } \Phi_M(\sigma) \leqslant \phi_i(f_n)\} \\ &\leqslant \Psi(f_n, \max\{\phi_i(f_n), \mu_i(f_n)\}). \quad \blacksquare\end{aligned}$$

As our final result of this section we prove an analog of the Blum speed-up theorem.

THEOREM 10. *For every inference complexity measure $\{\Phi_i\}$ and for every total limiting recursive functional ψ: $\mathscr{R} \times N \to N$ there exists a set $\mathscr{S}_\psi \subseteq \mathscr{S}_*$ of total recursive functions such that $\mathscr{S}_\psi \in EX$ and for all M, if $\mathscr{S}_\psi \subseteq EX(M)$, then there exists an M' such that $\mathscr{S}_\psi \subseteq EX(M')$ and $(\forall^\infty f \in \mathscr{S}_\psi)[\psi(f, \Phi_{M'}(f)) \leqslant \Phi_M(f)]$.*

Proof. The proof will parallel the approach of Blum and will be a modification of the construction given in Theorem 8. However, instead of constructing each f_n independently, each f_n will depend on f_m for $m \leqslant n$. Also, in the manner of Blum, the diagonalization against M will be more severe than that against $M + 1$.

Let ψ be a total limiting recursive functional. Without loss of generality we can assume that $\psi(f, 0) \geqslant f(0)$, that $\psi(f, n) \geqslant \mu(\psi, f, n)$ $\mu(\psi, f, n)$ is the point of convergence of ψ on inputs f and n), and that $\psi(f, n+1) > \psi(f, n)$ for all $f \in \mathscr{R}$, since otherwise we could easily define a limiting recursive functional ψ' with these properties and such that $\psi(f, n) \leqslant \psi'(f, n)$. We first define a total limiting recursive functional Δ, which is similar to that of Theorem 9, as follows:

$$\Delta(\sigma, s) = \sum_{M \leqslant \sigma(0)} \max\{s, \Phi_i(\sigma') \mid \sigma' \subseteq \sigma \text{ and } i \leqslant s \text{ and } \delta_i(\sigma') \leqslant \Omega(\sigma, s)\},$$

where Ω is defined in exactly the same way as Ψ was in Theorem 9 above. As above it is easy to see that Δ is a total limiting recursive functional. Next, let $\phi_i = \psi$ and define the total limiting recursive functional Ψ as follows:

$$\Psi(f, 0) = \phi_i(f, 0)$$
$$\Psi(f, n+1) = \phi_i(f, \Delta(f, \Psi(f, n))).$$

We now define as in the construction given in Theorem 8 a collection of functions of finite support $\phi_{\alpha(n)}$ in such a way that f_n will be the 0-completion of $\phi_{\alpha(n)}$. The construction of $\phi_{\alpha(n)}$ employs n movable markers, one for each $M \leqslant n$, whose position at stage s is given by $a^s_{n,M}$. Initially, $a^0_{n,M} = M$. The construction will be so arranged that for all stages s and for all $M \leqslant n$, $a^s_{n,M} \bmod n \equiv M$. We use $\sigma^s_{n,M}$ to denote the largest initial segment of $\phi^s_{\alpha(n)} \cup \{(a^s_{n,m}, 0) \mid M < m \leqslant n\}$, and $x^s_{n,M}$ to denote the first integer $x > a^s_{n,M}$ such that $x \notin \operatorname{dom} \phi^s_{\alpha(n)}$ and $x \equiv M \bmod n$. We fix some computational complexity measure $\{\tilde{\Phi}_i\}$. Also, for each $M \leqslant n$, we use $p^s_{n,M}$ to denote $\phi_M(\sigma)$, where σ is the largest initial segment of $\phi^s_{\alpha(n)} \cup \{(a^s_{n,m}, 0) \mid m \leqslant n\}$ such that $\tilde{\Phi}_M(\sigma) \leqslant s$. We will also assume that stage s in the construction of $\phi_{\alpha(n)}$ is carried out after stage s in the construction of $\phi_{\alpha(k)}$ for all $k < n$. The markings **a**, **b** and **c** have the same meaning below as above. By a straightforward application of the recursion theorem we can arrage that $\phi_{\alpha(n)}(0) = \langle n, \alpha(n) \rangle$, but we will omit the formal details of this application so as not to obscure the essential ideas.

Stage s. (of $\phi_{\alpha(n)}$)
Do the following four steps in order.

(1) If there is some $m \leqslant n$ such that $\Psi(\sigma^s_{n,m}, m) \neq \Psi(\sigma^{s-1}_{n,m}, m)$ or $\Psi(\sigma^s_{n,m}, m) \neq \Psi(\phi^s_{\alpha(n)} \cup \{(a^s_{n,M}, 0) \mid M \leqslant n\}, m)$, then for the least such m

(a) for all $M \leqslant n - m$, define $a^{s+1}_{n,M} = x^s_{n,M}$, and

(b) for all $M \leqslant n - m$, mark M **c**, and

(c) define $\phi^{s+1}_{\alpha(n)} = \phi^{s}_{\alpha(n)} \cup \{(a^{s}_{n,M}, 0) \mid M \leqslant n-m\}$, and

(d) go to stage $s+1$.

(2) else for each $M \leqslant n$

(a) if M is marked **b** and $\Phi_M(\sigma^s_{n,M}) \leqslant \Psi(\sigma^s_{n,M}, n-m)$, then mark M **c**.

(b) if M is marked **c** and $\Phi_M(\sigma^s_{n,M}) > \Psi(\sigma^s_{n,M}, n-m)$, then mark M **b**.

(3) Set $C = \phi$.
For each $M \leqslant n$

(a) if M is marked **c** and for all $k < n$, M is marked **b** at stage s of $\phi_{\alpha(k)}$ and $\tilde{\Phi}_{p^s_{n,M}}(a^s_{n,M}) \leqslant s$, then

(i) mark M **a**, and

(ii) define $a^{s+1}_{n,M} = x^s_{n,M}$, and

(iii) set $C = C \cup \{(a^s_{n,M}, 1 \dot{-} \phi_{p^s_{n,M}}(a^s_{n,M}))\}$,

(b) else if M is marked **c** and for all $k < n$, M is marked **b** at stage s of $\phi_{\alpha(k)}$ and $p^s_{n,M} \neq p^{s-1}_{n,M}$, then

(i) mark M **c**, and

(ii) define $a^{s+1}_{n,M} = x^s_{n,M}$, and

(iii) set $C = C \cup \{(a^s_{n,M}, 0)\}$,

(c) else if M is marked **a** or **c** and for all $k < n$, M is marked **b** at stage s of $\phi_{\alpha(k)}$, then

(i) define $a^{s+1}_{n,M} = a^s_{n,M}$, and

(ii) set $C = C \cup \{(x^s_{n,M}, 0)\}$,

(d) else

(i) define $a^{s+1}_{n,M} = x^s_{n,M}$, and

(ii) set $C = C \cup \{(a^s_{n,M}, 0)\}$.

(4) Define $\phi^{s+1}_{\alpha(n)} = \phi^s_{\alpha(n)} \cup C$, and go to stage $s+1$. *End stage s.*

Let f_n be the 0-completion of $\phi_{\alpha(n)}$, for all n. As before each M in the construction of $\phi_{\alpha(n)}$ is eventually marked **a**, or **b** or **c**. If M is permanently marked **b**, then $\lim_{s \to \infty} a^s_{n,M} = \infty$ by step 3d. Let $k \leqslant M$ be the least integer such that M is permanently marked **a** or **c** for $\phi_{\alpha(k)}$. Then by step 3c we have $\phi_M(f_k)\downarrow$ and $\lim_{s \to \infty} a^s_{k,M} = a_{k,M}$ exists. But, then by the argument in the proof of Theorem 8, for this k if M is permanently marked **a** for f_k then $\phi_{\phi_M(f_k)}(a_{k,M}) \neq f_k(a_{k,M})$, and if M is permanently marked **c** for f_k then

$\phi_{\phi_M(f_k)}(a_{k,M})\uparrow$, so in either case $f_k \notin EX(M)$ and $\mathscr{S}_\psi \not\subseteq EX(M)$. Also, by step 3d for all $n > k$ such that M is permanently marked **a** or **c** we have $\lim_{s\to\infty} a^s_{n,M} = \infty$. On the other hand, if there is no such k then M must be permanently marked **b** for all f_n, and so $\Phi_M(f_n) > \Psi(f_n, n-M)$ for all $n \geqslant M$. Consequently, $\phi_{\alpha(n)}(x)\uparrow$ for at most one x and at most one n because of any given M. In other words, each M is responsible for at most one undefined value among all $\phi_{\alpha(n)}$. This information will be encoded as $\langle k, a_{k,M}\rangle$ if such a k exists, or as $\langle k, 0\rangle$ if no such k exists, and will be denoted by u_M.

We now show that for each M there exists an M' such that $S_\psi \subseteq EX(M')$ and for all sufficiently large k that $\Phi_{M'}(f_k) \leqslant \Delta(f_k, \Psi((f_k, k-M+1))$. Since $\Phi_M(f_k) \geqslant \Psi(f_k, k-M)$ and $\Psi(f_k, k-M) \leqslant \psi(f_k, \Delta(f_k, \Psi(f_k, k-M+1)))$, we will then have $\Phi_M(f_k) \geqslant \psi(f_k, \Phi_{M'}(f_k))$, which is the sought after speed-up. We first define the total recursive program transformation β as follows:

$$\phi_{\beta(v)}(\sigma) = \begin{cases} p_{\sigma,v}, & \text{if } \dim v < \pi_1(\sigma(0)) \text{ and } \sigma(\max_\sigma) \neq 0 \text{ and} \\ & \text{for some } k > \dim v,\ \max_\sigma \equiv k \bmod \pi_1(\sigma(0)), \\ \phi_{M_*}(\sigma), & \text{if } \dim v \geqslant \pi_1(\sigma(0)), \\ \phi_{\beta(v)}(\sigma - 1), & \text{otherwise,} \end{cases}$$

where the parameter v is a tuple of values which contains the information required to form the 0-completion of the desired function, and $\dim v$ denotes the number of values in the tuple v. Next we define the programs $p_{\sigma,v}$ by

$$\phi_{p_{\sigma,v}}(x) = \begin{cases} \sigma(x), & \text{if } \ x \in \operatorname{dom} \sigma, \\ 0, & \text{if } \ (\exists u_1,\dots, u_M)[v = \langle u_1,\dots, u_M\rangle] \text{ and } x > 0 \\ & \text{and } (\exists j \leqslant M)[\langle \pi_1(\sigma(0)), x\rangle = u_j], \\ \phi_{\pi_2(\sigma(0))}(x), & \text{otherwise.} \end{cases}$$

We first suppose that $k > M$ and let $x_k = \max\{x \mid f_k(x) \neq 0 \text{ and } x \equiv j \bmod k$ for some $M < j \leqslant k\}$. Then $\sigma_k = f_k \mid x_k$ will contain the completed value of all those x such that $\phi_{\alpha(k)}(x)\uparrow$ because of some machine j, where $M < j \leqslant k$. By our remarks above concerning the information u_i we see that $v = \langle u_1,\dots, u_M\rangle$ contains the completed value of all those x such that $\phi_{\alpha(k)}(x)\uparrow$ because of some machine j, where $j \leqslant M$. Therefore, $\phi_{p_{\sigma_k,v}}$ is the 0-completion of $\phi_{\alpha(k)}$, i.e., $\phi_{p_{\sigma_k,v}} = f_k$ (observe that in the definition of $p_{\sigma_k,v}$ that $\pi_1(\sigma_k(0)) = k$ and $\pi_2(\sigma_k(0)) = \alpha(k)$). Letting $M' = \beta(v)$ and observing that dim $v = M < k$ we see from the definition of $\beta(v)$ that $\phi_{M'}(f_k) = p_{\sigma_k, v}$, so that $f_k \in EX(M')$. Since ψ is strictly increasing (in its second argument) and Δ and Ψ are clearly increasing, we have that for sufficiently large values of k that $\Psi(f_k, k - M + 1) \geqslant \beta(v)$. By an argument similar to that used in Theorem 9 above we see that $\delta_{\beta(v)}(f_k) \leqslant \Omega(f_k, \Psi(f_k, k - M + 1))$, and hence $\Phi_{M'}(f_k) \leqslant \Delta(f_k, \Psi(f_k, k - M + 1))$. Finally, supposing that $k \leqslant M =$ dim v we see from the definition of $\beta(v)$ that $\phi_{M'}(\sigma_k) = \phi_{M_*}(\sigma_k)$, and since $f_k \in \mathscr{S}_*$, we have $f_k \in EX(M')$, and therefore $\mathscr{S}_\psi \subseteq EX(M')$. ∎

We mention here also that Shäfer-Richter (1984) has shown that the Gap theorem does not hold for inference complexity measures by exploiting the fact that the number of mind changes forms a hierarchy, i.e., any increase in the number of mind changes results in added inference power and hence a new class of inferrable functions.

4. Further Considerations

In this section we briefly examine the influence which the order of presentation has on the complexity of inference. We will also consider extensions to other notions of inductive inference of our axiomatic approach to the complexity of inference. Changing the order of presentation, of course, will have no effect on inferrability itself, but it can have an effect on the complexity of inference as we shall see below. We will restrict our attention here to effective presentations of the input function. Let $\mathscr{RP}$ denote the set of recursive permutations. If $g \in \mathscr{RP}$, then we denote by $f\colon g$ the *g-presentation* of f, and $f\colon g \mid n$ will denote the initial segment $\langle (g(0), f(g(0))), ..., (g(n), f(g(n))) \rangle$ of this presentation. Similarly, $\tau\colon g$ will denote the g-presentation of the finite function τ, i.e., $\langle (g(i_1), \tau(g(i_1))), ..., (g(i_n), \tau(g(i_n))) \rangle$ where dom $\tau = \{g(i_1), ..., g(i_n)\}$ and $i_1 < i_2 < \cdots < i_n$. Also, $\phi_M(f\colon g) = \lim_{n \to \infty} \phi_M(f\colon g \mid n)$, and whenever $\phi_M(f\colon g)\downarrow$, $\mu_M(f\colon g) = \min\{n \mid (\forall m \geqslant n)[\phi_M(f\colon g \mid m) = \phi_M(f\colon g \mid n)]\}$. We denote by $f\colon g \in EX(M)$ that $\phi_M(f\colon g) = p$ and $\phi_p = f$.

We begin with an example involving the functions of finite support $\mathscr{S}_*$, the natural IIM M_* for $\mathscr{S}_*$, and the modulus complexity measure, which illustrates the dramatic increases and reductions which are possible when the order of presentation is changed. Here we will suppose that M_* has been extended to arbitrary enumerations in the obvious way so that $f\colon g \in EX(M_*)$ for all $f \in \mathscr{S}_*$ and all $g \in \mathscr{RP}$.

THEOREM 11. *For every $h \in \mathscr{RP}$ and for every IIM M such that $\mathscr{S}_* \subseteq EX(M)$,*

(1) $(\exists^{\infty} g \in \mathscr{RP})$ *such that*

(a) $(\exists^{\infty} f \in \mathscr{S}_*)[\mu_M(f: g) > h(\mu_{M_*}(f))]$,

(b) $(\exists^{\infty} f \in \mathscr{S}_*)[\mu_M(f) > h(\mu_{M_*}(f: g))]$,

(2) $(\exists^{\infty} f \in \mathscr{S}_*)$ *such that*

(a) $(\exists^{\infty} g \in \mathscr{RP})[\mu_M(f: g) > h(\mu_{M_*}(f))]$,

(b) $(\exists^{\infty} g \in \mathscr{RP})[\mu_M(f) > h(\mu_{M_*}(f: g))]$.

Proof. We will give the construction only for part 1 since the construction for part 2 is similar. We will define a $g \in \mathscr{RP}$ which satisfies parts 1a and 1b, and whose values on the odd integers are irrelevant so that there will exist infinitely many such g. The subset of $\mathscr{S}_*$ which we use consists of all functions f such that $f =^1 0$ and $f(x) = 0$ for all odd x, and $f(x) \leqslant 2$ for all x. The function g is an interchange function on the even integers and is defined by

for even x,

$$g(x) = \begin{cases} \min\{y \mid y > h(x) \text{ and } y \text{ is even and } y \neq g(z) \text{ for any even } z < y\}, \\ \qquad \text{if } x \neq g(z) \text{ for any } z < x, \\ z, \qquad \text{where } z = \min\{y \mid g(y) = x\}, \text{ otherwise,} \end{cases}$$

for odd x,

$$g(x) = 2 \times r(\lceil x/2 \rceil) + 1,$$

where $r \in \mathscr{RP}$ is an arbitrary recursive permutation. For any even integer x we define

$$f_{x,1}(y) = \begin{cases} 1, & \text{if } \; y = x, \\ 0, & \text{otherwise.} \end{cases}$$

$$f_{x,2}(y) = \begin{cases} 2, & \text{if } \; y = x, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, for all even x $\mu_{M_*}(f_{x,1}) = \mu_{M_*}(f_{x,2}) = x$. Suppose for all but finitely many even x that $f_{x,1}: g \in EX(M)$ and $f_{x,2}: g \in EX(M)$, since otherwise part 1a is immediate. Then, $\phi_M(f_{x,1}: g)\downarrow$ and $\phi_M(f_{x,2}: g)\downarrow$. Since $f_{x,1}: g \mid n = f_{x,2}: g \mid n$ for all $n < g(x)$, it is clear that either $\mu_M(f_{x,1}: g) \geqslant g(x) > h(\mu_{M_*}(f_{x,1}))$ or $\mu_M(f_{x,2}: g) \geqslant g(x) > h(\mu_{M_*}(f_{x,2}))$. Therefore, $(\exists^{\infty} f \in \mathscr{S}_*)$ such that $\mu_M(f: g) > h(\mu_{M_*}(f))$. Similarly, since $g(g(x)) = x$, $\mu_{M_*}(f_{g(x),1}: g) = \mu_{M_*}(f_{g(x),2}: g) = x$, but again either $\mu_M(f_{g(x),1}) > h(\mu_{M_*}(f_{g(x),1}))$ or $\mu_M(f_{g(x),2}) > h(\mu_{M_*}(f_{g(x),2}))$, and $(\exists^{\infty} f \in \mathscr{S}_*)$ such that $\mu_M(f) > h(\mu_{M_*}(f: g))$. ∎

Thus, we see that the inference complexity can be reduced as well as increased by changing the order of presentation. We point out that Theorem 11 remains valid if the modulus complexity measure is replaced by any inference complexity measure which is recursively related to it, i.e., $\mu_M \leqslant \Phi_M \leqslant r \circ \mu_M$ for some $r \in \mathscr{R}$. Moreover, parts 1a and 2a remain valid for any inference complexity measure for which $\Phi_M \geqslant \mu_M$. However, for parts 1b and 2b this is not so, since we can define the following inference complexity measure,

$$\hat{\Phi}_i(\tau) = \max\{\mu_i(\tau), x \mid (x, y) \in \tau \text{ for some } y\}.$$

Theorem 11 shows that there can in general be no recursive bound for the increase (or decrease) in inference complexity when the order of presentation is changed. We now show that a limiting recursive upper bound does exist.

THEOREM 12. *For every inference complexity measure* $\{\Phi_i\}$ *there exists a total limiting recursive functional* Ψ *such that for all* M *there exists an* M' *such that for all* $g \in \mathscr{RP}$ *and for all* $f \in \operatorname{dom} \phi_M$, $\phi_{M'}(f\colon g) = \phi_M(f)$ *and* $\Phi_{M'}(f\colon g) \leqslant (f, g, \Phi_M(f))$.

Proof. Define M' as follows:

$$\phi_{M'}(\tau) = \begin{cases} \phi_M(\sigma'), & \text{where } \sigma' \subseteq \tau \text{ is the largest initial} \\ & \text{segment such that } \phi_M(\sigma') \neq \phi_M(\sigma' - 1) \\ & \text{and } (\forall \sigma'' \subseteq \sigma')[\Phi_M(\sigma'') \leqslant \operatorname{card}(\tau)], \\ \phi_M(\phi), & \text{if no such } \sigma' \text{ exists.} \end{cases}$$

Thus, M' simply ignores any information contained in the input segment τ which is not arranged as an initial segment. Observe that if $\phi_M(\phi)\downarrow$ then $\phi_{M'}$ is total, and if $\phi_M(f)\downarrow$, then $\lim_{n \to \infty} \phi_{M'}(f\colon g \mid n) = \phi_M(f)$ even if $\langle \phi_M(f \mid n) \rangle_{n \in N}$ is a finite sequence. We now define Ψ by

$$\Psi(\sigma, g, z) = \begin{cases} \Phi_{M'}(\tau : g), & \text{where } \tau \text{ is the least finite function such that} \\ & \sigma' \subseteq \tau : g \subseteq \sigma \text{ and } \tilde{\Phi}_{M'}(\tau : g) \leqslant \max_{\sigma'} \text{ where} \\ & \sigma' \subseteq \sigma \text{ is the largest initial segment such that} \\ & \text{(i)} \quad (\forall \sigma'' \subseteq \sigma')[\tilde{\Phi}_M(\sigma'') \leqslant \max_\sigma], \\ & \text{(ii)} \quad \phi_M(\sigma') \neq \phi_M(\sigma' - 1), \\ & \text{(iii)} \quad \Phi_M(\sigma') \leqslant z, \\ \Phi_{M'}(\phi), & \text{if no such } \tau \text{ or } \sigma' \text{ exist.} \end{cases}$$

We first point out that Ψ is a limiting recursive functional in its argument, and a recursive functional in its second argument. Also, for any given $g \in \mathscr{RP}$ $\Psi(\cdot, g, \cdot)$ is total since there can be only finitely many choices of initial segments σ' with $\Phi_M(\sigma') \leqslant z$ and $\phi_M(\sigma') \neq \phi_M(\sigma' - 1)$ (recall we are assuming that Φ_M is an increasing functional), and therefore for any $g \in \mathscr{RP}$ there can be only finitely many $\tau : g$. It is clear for any $f \in \operatorname{dom} \phi_M$ and $g \in \mathscr{RP}$ from the definition of $\Psi(f, g, z)$ that in the limit $\sigma' = f \mid \mu_M(f)$ and $\Phi_M(\sigma') = \Phi_M(f)$. Also, from the definition of M' we see that $\mu_{M'}(f\colon g)$ is the least integer n such that $\sigma' \subseteq f\colon g \mid n$. Thus, $\Phi_{M'}(f\colon g) = \Phi_{M'}(f\colon g \mid \mu_{M'}(f\colon g)) = \Psi(f, g, \Phi_M(f))$.

There is another type of inductive inference called BC inference in Case and Smith (1983), where the IIM is not required to converge to a fixed program for the input function, but only required to produce a (possibly infinite) sequence of programs whose behavior converges to that of the input function. Thus the convergence is only second order. Barzdin (1974) also studied the notion of BC inference which he called GN^∞ inference. More formally, we say that an IIM M BC^n identifies (and write $f \in BC^n(M)$) if and only if there exists and integer k such that $(\forall m \geqslant k)[\phi_{\phi_M(f \mid m)} =^n f]$. We also define

$$BC^n = \{\mathscr{S} \mid (\exists M)[\mathscr{S} \subseteq BC^n(M)]\}.$$

Clearly, in the case BC inference the number of mind changes does not make any sense as a measure of complexity, but a (second-order) modulus function μ^2 does exist and can be used as a basis for defining an a.u.c. type complexity measure. We can define

$$\mu^2(f, k) = \min\{n \mid (\forall m \geqslant n)[\phi_{\phi_M(f \mid m)} =^k f]\}.$$

Observe that in contrast with EX inference, where $\phi_M(f)\downarrow$ and $f \notin EX^*(M)$ is possible, since the convergence is only second-order, the criterion for success must be a part of the notion of convergence. Thus, the intuitive notion for the complexity of BC inference is well founded and we could proceed to develop an axiomatization analagous to that for EX inference above. However, we instead conclude with a strengthening of Theorem 1, which follows from the construction of $f_{n,h,j}$ in the proof since the diagonalization there was against any program produced by any IIM $M \leqslant j$.

THEOREM 13. *There exists a* $g \in \mathscr{R}$ *and IIMs* $M_0, M_1, \ldots$ *such that for all* $h \in \mathscr{R}$ *and for all* $n \in N$ *there exists* $\mathscr{S}_{n,h} \in EX^n$ *and*

(1) *for all* $k < n$ $\mathscr{S}_{n,h} \subseteq EX^{n-k}(M_k)$

and

$$(\forall^\infty f \in \mathcal{S}_{n,h})\ [h(\Phi_{M_{k-1}}(f)) \leqslant \Phi_{M_k}(f) \leqslant g(h(\Phi_{M_{k-1}}(f)))],$$

(2) *for all M and for all $k \leqslant n$ and for all $m < k$ if $\mathcal{S}_{n,h} \subseteq BC^m(M)$ then*

$$(\forall f \in \mathcal{S}_{n,h})[\Phi_M(f) > h^{k-m}(\Phi_{M_m}(f))].$$

5. Conclusions and Open Problems

Area under the curve and abstract complexity measures were introduced for inductive inference. Accuracy verses complexity trade-off results were proven for the a.u.c. measures. The existence of arbitrarily difficult to infer sets of functions and analogues of the speed-up and compression theorems of (Blum, 1967) established. The above results used total limiting recursive functionals to name complexity of inference classes. It remains open to show similar results when arbitrary partial limiting recursive functionals are used as complexity class names. Most of our results relied on the increasing domain order enumeration of some function as input to an IIM. Do the same results hold when the input is given in an arbitrary order? It would also be interesting to develop a notion of relative inferribility, where one IIM had available to it the guesses made by some other IIM, on the same input, without incurring the addition cost of simulation. One approach to relative inferribility is to use the enumeration reducibility of Selman (1978).

Acknowledgments

We are grateful to the Departments of Computer Science at the Universities of Pittsburgh, Maryland and Purdue for providing computer time. CSNET was instrumental in the writing of this paper. Special thanks go to our colleague Paul Young for informative conversations. The referee made some suggestions which were valuable in improving the final form of this paper.

Received August 4, 1983; accepted May 21, 1984

References

Angluin, D. (1978), On the complexity of minimum inference of regular sets, *Inform. and Control* **39**, 337–350.

Angluin, D., and Smith, C. (1983), A survey of inductive inference: theory and methods, *Comput. Surveys* **15**, 237–269.

BARZDIN, J. (1974), Two theorems on the limiting synthesis of functions, *in* "Theory of Algorithms and Programs I," pp. 82–88, Latvian State University, Riga, U.S.S.R.

BARZDIN, J., AND FREIVALDS, R. (1972), On the prediction of general recursive functions, *Soviet Math. Dokl.* **13**, 1224–1228.

BLUM, L., AND BLUM, M. (1975), Toward a mathematical theory of inductive inference, *Inform. and Control* **28**, 125–155.

BLUM, M. (1967), A machine-independent theory of the complexity of recursive functions, *J. Assoc. Comput. Mach.* **14**, 322–326.

CASE, J., AND NGO-MANGUELLE, S. (1979), "Refinements of inductive inference by Popperian machines," Technical report, Depart. of Comput. Sci. State Univ. New York, Buffalo.

CASE, J., AND SMITH, C. (1983), Comparison of identification criteria for machine inductive inference, *Theoret. Comput. Sci.* **25**, 193–220.

FREIVALDS, R. (1975), On the complexity and optimality of computation in the limit, *in* "Theory of Algorithms and Programs, II," pp. 155–173, Lativan State University, Riga, U.S.S.R.

GOLD, E. (1967), Language identification in the limit, *Inform. and Control* **10**, 447–474.

GOLD, E. (1978), Complexity of automaton identification from given data, *Inform. and Control* **37**, 302–320.

GOLD, E. (1965), Limiting Recursion, *J. Symbolic Logic* **30**, 28–48.

HEMPEL, C. (1965), "Aspects of Scientific Explanation," The Free Press, New York.

KHODZHAYEV, J. (1970), "On the complexity of computation on Turing machines with oracles," Ph. D. dissertation, Tashkent.

KLETTE, R., AND WIEHAGEN, R. (1980), Research in the theory of inductive inference by GDR mathematicians—a survey, *Inform. Sci.* **22**, 149–169.

LYNCH, N., MEYER, A., AND FISCHER, M. (1976), Relativization of the theory of computational complexity, *Trans. Amer. Math. Soc.* **220**, 243–287.

MACHTEY, M., AND YOUNG, P. (1978), "An Introduction to the General Theory of Algorithms," North-Holland, New York.

PUDLAK, P. (1975), Polynomially complete problems in the logic of automated discovery, *in* "Lecture Notes in Comput. Sci." Vol. 32, pp. 358–361, Springer-Verlag, New York/Berlin.

PUDLAK, P., AND SPRINGSTEEL, F. (1979), Complexity in mechanized hypothesis formation, *Theoret. Comput. Sci.* **8**, 203–225.

PUTMAN, H. (1975), Probability and confirmation, *in* "Mathematics, Matter, and Method," Cambridge Univ. Press. Originally appeared in 1963 as a Voice of America Lecture.

ROGERS, H. (1958), Gödel numberings of partial recursive functions, *J. Symbolic Logic* **23**, 331–341.

ROGERS, H. (1967), "Theory of Recursive Functions and Effective Computability," McGraw Hill, New York.

SELMAN, A. (1978), Polynomial time enumeration reducibility, *SIAM J. Comput.* **7**, 440–457.

SHÄFER-RICHTER, G. (1984), Über Eingabeabhängigkeit und Komplexität von Inferenzstrategien, Diplom-Mathematikerin, Technische Hochschule, Aachen.

SHOENFIELD, J. (1971), "Degrees of Unsolvability," North-Holland, Amsterdam.

SYMES, D. (1971), "The extension of machien-independent computational complexity theory to oracle machine computation and to the computation of finite functions," Ph. D. dissertation, CSRR 2057, Univ. of Waterloo, Ontario, Canada.

YOUNG, P. (1971), Speed-ups by changing the order in which sets are enumerated, *Math. Systems Theory* **5**, 148–152.